DOCUMENT RESUME

ED 276 314 FL 016 296

AUTHOR Kotsinas, Ulla-Britt
TITLE On the Acquisition of Vocabulary in Immigrant Swedish.
PUB DATE 83
NOTE 28p.; In: Psycholinguistics and Foreign Language Learning. Papers from a conference (Stockholm, Sweden and Abo, Finland, October 25-26, 1982); see FL 016 291.
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Adult Learning; Difficulty Level; Foreign Countries; *Immigrants; Interpersonal Communication; *Learning Processes; Learning Strategies; Psycholinguistics; *Second Language Learning; *Swedish; Uncommonly Taught Languages; *Vocabulary Development; Word Frequency; Word Recognition

ABSTRACT

A study of the acquisition of Swedish vocabulary by adult immigrants examined both the kinds and contexts of words learned. It found that simple and common words form the bulk of the vocabulary learned, and that frequency of input, complexity of input, and communicative need guide the learning process. Frequency refers to the frequency of both target language words and words in situations where the learner usually interacts linguistically. Complexity may be the words' similarity to native language words, word length, semantic complexity, or degree of concreteness or abstractness. Communicative need might be for words connected to the situation in which the learner uses the target language, words expressing emotions and attitudes, or form words. The investigation found three strategies of word learning: the learning of whole "chunks," the learning of words occurring in certain contexts, and copying. (MSE)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED276314

ON THE ACQUISITION OF VOCABULARY IN IMMIGRANT SWEDISH

Ulla-Britt Kotsinas
Department of Scandinavian Languages

University of Stockholm

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

☒ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

N. Enkvist

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

FL 016 296

BEST COPY AVAILABLE

Ulla-Britt Kotsinas
Department of Scandinavian Languages,
University of Stockholm

ON THE ACQUISITION OF VOCABULARY IN IMMIGRANT SWEDISH

Introduction

What kind of words does a learner pick up from the flow of speech around him when he learns a language solely by listening to target language speakers? Do different learners pick up roughly the same lexical items, or does every learner haphazardly learn words quite differently from other learners? Does the learner acquire words from all word-classes, or does he concentrate on learning content words only? And is it possible to draw any conclusions about the process of language learning from the learners' spontaneous utterances?

The following paper is a report concerning the cocabulary of six immigrants in Sweden. I shall account for what types of words they have learnt during their stay in the country, and shall also suggest a few factors that seem to operate during the process of acquisition of vocabulary.

The material

The material for this study consists of 22 hours of tape-recordings made on 15 different occasions during the first third of a 240 hour course in Swedish. This course was taught by the author.

All through the course the tape-recorder was on and registered all speech, spontaneous as well as classroom oriented. Only material from the first third of the course has been used in the study, and a total of approximately 23,000 words have been transcribed. From these were deleted all words uttered by the teacher, and all words related to the teaching situation. Thus, only

sequences of informal conversations are used. The remaining 10,430 text-words (tokens) constitute the material for this study.

The informants

The informants for this study were four Greek men, a Greek woman and a Polish woman. Their ages ranged from 26 to 49, and all had lived in Sweden for between 3-15 years (table 1). All informants worked as cleaning personnel at a big office in Stockholm. With the exception of the Polish woman, none of them had received any education beyond grammar school. None of them had attended any courses in Swedish earlier, or otherwise learnt any language other than their mother tongue.

Table 1. Informants

Georgios (Ge)	Nat.	Age	Years in Sweden	Education	Lexical tokens	Lexical types
Georgios (Ge)	Gr.	43	15	6 yrs	812	196
Dimitris (Di)	Gr.	44	9	6 yrs	4329	381
Takis (Ta)	Gr.	29	6	6 yrs	2404	318
Katarina (Ka)	Gr.	40	8	6 yrs	237	101
Nikolaos (Ni)	Gr.	26	3	6 yrs	230	94
Maria (Ma)	Po.	49	3	12 yrs	2418	360

The Greek informants lived rather isolated from Swedish society, viz. they were married to compatriots and spoke Greek at home and with their friends. At work they had a minimum of contact with their Swedish supervisors and only occasionally exchanged a few words with the rest of the staff. For necessary contacts with Swedish authorities, hospitals, etc., they used municipal interpreters. They all planned to return to Greece some day.

The Polish woman, Maria, was married to a Swede and her contacts with Swedish society were much closer than those of the Greek informants. She spoke Swedish with her husband and his relatives, and almost never met with compatriots in Sweden. She did not plan on returning to Poland.

Immigrant Swedish

The term Immigrant Swedish (abbr. IS) stands for the spoken Swedish used by the six informants. The term Native Swedish (abbr. NS) is used for all kinds of spoken Swedish, and is primarily based upon an investigation of written and spoken Swedish (Einarsson 1978).

Here follow a few examples of Immigrant Swedish. The transcriptions in Swedish are followed by literal translations to English, these being as true to the Swedish text as possible.

(1) Ge: vi alla samma förstår/ vi inte så bra svenska eller Takis inte lära eller Dimitris eller Katarina/ vi lära samma alla förstår

(We all same, understand. We not so good Swedish or Takis not learn or Dimitris or Katarina. We learn all same, understand?)

(2) Di: de kanske fem fem timmar normal// den inte jobba// bara i i skolan// läsa på hemma// kommer i skola de bra// den jobba klockan sex/ jobba kommer på skola// inte läsa ingenting// ja? – – – på kvällen arbeta// på dagar arbeta// kommer skola sex timma

(It maybe five five hour normal. It not work. Only in in school. Read on home, come in school it good. It work six o'clock, work come on school. Not read nothing. Yes? – – – On evening work, on days work. Come school six hour.)

(3) Ta: den chilenare pojk fråga mej ja inte förklara om inte/ riktigt bra förklara – – – den chilenare pojk mycke bättre// han läsa varje dag på tre timma

(It Chilean boy ask me I not explain, if not real good explain. It Chilean boy much better. He read every day on three hour.)

(4) Ka: fel/ Ulla// betala// varje månad fel fel fel/ skriva

(Wrong, Ulla, pay. Every month wrong wrong wrong, write.)

(5) Ni: ja ja tycker// ja tycker att svårt svenska

(Yes I think, I think that difficult Swedish.)

(6) Ma: nej vi vi blir trött trött helt enkelt/ de går inte på oss/ nej// nej// Lära sej sakta men bra// de e bra// säkert// vi kan inte lära så fort svenska språk// vi kan inte

(No, we we get tired tired quite simply. It is not possible on us. No. No. Learn slowly but good. It is good. Sure. We can not learn so fast Swedish language. We can not.)

In example (1) Georgios explains that no one in the group knows Swedish very well and that they all have to learn in the same way. In example (2) Dimitris expresses the opinion that it is not possible to study Swedish if one has to work at the same time, because one is tired after work. In example (3) Takis tells us that he once had to explain something to another immigrant, and that he could not do it. The other boy knew Swedish much better because he studied Swedish every day. In example (4) Katarina complains about her salary. The employer always pays her the wrong sum. In example (5) Nikolaos states that Swedish is difficult and in (6) Maria says that one can't study Swedish every day, because one gets tired and because it is impossible to learn Swedish fast.

It is evident even from these short excerpts of the material that the kind of Swedish these informants use deviates in many ways from ordinary Swedish as spoken by native Swedes. Indeed it is sometimes rather difficult even to understand the meaning of the utterances, especially if one is not accustomed to the language that these people use. In fact, some of the utterances are possible to interpret only if one compares the use of certain lexical items in the whole material. Some clues to the interpretation can also be found in many of the Greek utterances that sometimes accompany the attempts in Swedish.

Number of words

As can be seen from table 1, one of the informants, Dimitris, spoke a lot during the conversations, while others (e.g. Katarina and Nikolaos) participated much less frequently. All together, Dimitris uttered 4329 words (tokens), Takis and Maria about 2400 words each and Katarina and Nikolaos only about 230 words each. Georgios, who was present at only one of the 15

analysed conversations, uttered 812 words. Table 1 also shows that some of the informants used a larger vocabulary than others. Dimitris used 381 different words (types) and Katarina and Nikolaos used only approximately 100 different words each. One should note, however, that the difference between Dimitris and Maria or between Dimitris and Takis regarding number of types is much smaller than the difference between the same informants regarding the number of tokens. Although Dimitris was the dominating speaker during the conversations, his vocabulary was very small and he had to use the few words he knew repeatedly in different contexts. The assumption that Dimitris used nearly all the Swedish words he knew to express his communicative intentions cannot be proven here, but it can be inferred. Judging from my experience as the teacher of the group, I also believe that Maria had a greater vocabulary than she demonstrated during the conversations. Georgios and Takis might also have had a slightly larger vocabulary than they actively used, while on the other hand Katarina and Nikolaos, like Dimitris, probably used all the Swedish words they knew over and over again.

Number of tokens

Table 2 below displays the number of tokens of each word-class for each informant and for the whole group. The percentages for each word-class are also shown. The far right hand column of this table contains the percentages of the same wordclasses in spontaneous spoken Swedish as used by native speakers (Einarsson 1978:166).

Looking at the two final columns in table 2 (those labelled Tot. and SS), it is obvious that there are differences regarding the distribution of word-classes. In Immigrant Swedish, nouns and verbs are used almost twice as frequently as they are in Native Swedish, and interjections and numerals have an even higher frequency ratio when compared to the Swedish material. Conversely, auxiliaries hardly exist at all in IS, and adverbs, pronouns and conjunctions have a considerably lower percentage in IS than in NS.

If we now consider the individual informants, we notice that Dimitris and Katarina use more nouns than the others. Almost 30% of Katarina's words are nouns, for example. Georgios has the highest percentage of verbs, almost

Table 2. Number of tokens

	Ge	Di	Ta	Ka	Ni	Ma	Tot.	SS
nouns	143(17.6)	955(22.1)	455(18.9)	70(29.5)	42(18.3)	292(12.1)	1957(18.8)	(9.9)
verbs	160(19.7)	714(16.5)	406(16.9)	21(8.9)	41(17.8)	344(14.2)	1686(16.2)	(9.3)
aux.	22(2.7)	28(0.6)	37(1.5)	5(2.1)	1(0.4)	43(1.8)	136(1.3)	(11.0)
adj.	21(2.6)	182(4.2)	64(2.7)	23(9.7)	11(4.8)	135(5.6)	436(4.2)	(4.6)
adv.	101(12.4)	623(14.4)	223(9.3)	46(19.4)	26(11.3)	411(17.0)	1430(13.7)	(18.8)
pron.	155(19.1)	738(17.0)	439(18.3)	24(10.1)	35(15.2)	526(21.7)	1917(18.4)	(23.6)
interj	23(2.8)	210(4.8)	237(9.8)	14(5.9)	25(10.9)	288(11.9)	797(7.6)	(2.2)
num.	23(2.8)	196(4.5)	104(4.3)	5(2.1)	11(4.8)	45(1.9)	384(3.7)	(0.3)
prep.	58(7.1)	259(6.0)	194(8.1)	13(5.5)	14(6.1)	96(4.0)	634(6.1)	5.3)
i.art.	11(1.3)	79(1.8)	33(1.4)	1(0.4)	1(0.4)	17(0.7)	142(1.4)	(1.3)
conj.	61(7.5)	75(1.7)	104(4.3)	3(1.3)	4(1.7)	125(5.2)	372(3.6)	(8.7)
other	34(4.1)	270(6.3)	108(4.5)	12(5.1)	19(8.3)	96(3.9)	539(5.1)	
Tot.	812(100)	4329(100)	2404(100)	237(100)	230(100)	2418(100)	10430(100)	(100)

20%. Except for a very high percentage of interjections (11.9%), Maria seems to be the informant whose numbers are closest to the Swedish percentage numbers.

Thus, despite differences between IS and NS, it is apparent that all informants, even those who speak very little Swedish, use words from all word-classes.

In the table, the group 'Other' consists of interrupted or uninterpretable words.

Number of types

From table 3 below one can see how many different words (types) each informant uses within each word-class, and the percentages for each word-class in the vocabulary. Numbers to the far right represent total occurrences for the whole group. (Since many words are used by more than one informant, these figures cannot be the sum of the figures calculated for each informant.) There are no corresponding Swedish figures to compare with, but it seems quite evident that the vocabulary of all the informants is very meager. We see, for example that the largest word-class for all the informants except Nikolaos is the noun class. For many of the informants, nouns are used twice or even three times as much as verbs. Also, even though content words are many (as is the case in all natural languages) there are a number of function words in the material. Prepositions, for instance, are represented by at least three different words in the speech of all informants, even those who have a very small vocabulary.

One probable explanation regarding the fact that content words are represented in a much higher percentage in Immigrant Swedish than in Native Swedish is that the immigrant repeats himself quite frequently. (For a discussion of this, see Kotsinas 1982a or 1983b.)

Table 3. Number of types

	Ge	Di	Ta	Ka	Ni	Na	Tot.
Nouns	50(25.5)	157(41.2)	116(36.5)	28(27.7)	21(22.3)	103(28.6)	275(38.2)
verbs	38(19.4)	51(13.4)	59(18.5)	17(16.8)	24(25.5)	59(16.4)	114(15.8)
aux.	5(2.5)	5(1.3)	5(1.6)	1(1.0)	1(1.1)	5(1.4)	8(1.1)
adj.	12(6.1)	30(7.9)	19(6.0)	12(11.9)	8(8.5)	41(11.4)	76(10.6)
adv.	24(12.2)	30(7.9)	34(10.7)	15(14.8)	11(11.7)	58(16.1)	82(11.4)
pron.	31(15.8)	24(6.3)	31(9.7)	9(8.9)	9(9.6)	40(11.1)	50(6.9)
interj.	7(3.6)	23(6.0)	10(3.1)	7(6.9)	5(5.3)	14(3.9)	28(3.9)
num.	4(7.1)	51(13.4)	30(9.4)	5(4.9)	9(9.6)	21(5.8)	63(8.8)
prep.	8(4.1)	4(1.0)	8(2.5)	4(4.0)	3(3.2)	10(2.8)	12(1.7)
ind.art.	1(0.5)	2(0.5)	1(0.3)	1(1.0)	1(1.1)	2(0.5)	2(0.3)
conj.	6(3.1)	4(1.0)	5(1.6)	2(1.9)	2(2.1)	7(2.0)	9(1.2)
tot.	196(100)	381(100)	318(100)	101(100)	94(100)	360(100)	719(100)

Frequently occurring words

Words can occur frequently in a material like this for two reasons: either they are used by all or almost all the informants, or are used in many tokens (by all or only a few informants). Table 4 shows the 100 most frequent words in the material with regard to the number of tokens. It is evident that the most frequently occurring words are also those used by most of the informants.

Of the 25 most frequently occurring words in the list, 23 are used by 5 or 6 informants. Only 2 are used by 4 speakers, and no word is used by less than 4 speakers. Words used in less than 75 tokens in the whole material (rank nr 25 and higher) seem to be more likely to be used by less than 6 informants, and the less frequent a word is, the more possible it seems that it is not used by all informants. Words used in only a few tokens are used mostly by one single informant.

There are no word frequency lists for spoken Swedish to compare this material to. Instead, I have made a comparison between my material and a frequency dictionary of written Swedish, the so called Tiotusen i topp (Top Ten Thousand, Allén 1972, abbr. TTT). This dictionary is based on a comprehensive investigation of present-day Swedish newspaper texts (about one million words). Most of the texts are reports from foreign correspondents, articles on cultural subjects and general news reports.

Table 4. Top 100 in Immigrant Swedish

Rank order	Nr of inf.	Nr of tokens	Rank order	Nr of inf.	Nr of tokens
1. inte (not)	6	459	11. mycket (very)	6	124
2. det (it, neutr.g.)	5	413	12. men (but)	5	119
3. på (on)	6	332	13. i (in)	5	116
4. ja (yes)	6	325	14. du (you)	5	111
5. jag (I)	6	325	15. och (and)	5	108
6. den (it)	4	281	16. kanske (maybe)	4	103
7. komma (come)	6	234	17. bra (good)	6	101
8. nej (no)	6	186	18. förstå (understand)	6	100
9. vara (be)	5	177	19. varför (why)	6	94
10. en (a, one)	6	142	20. med (with)	6	91

Rank order		Nr of inf.	Nr of tokens
21.	lite (little)	6	82
	också (too)	5	82
23.	vad (what)	5	80
24.	eller (or)	5	76
25.	va (what? interj.)	5	74
26.	veta (know)	6	73
	han (he)	5	73
28.	kallas (be called)	5	71
29.	timme (hour)	4	67
30.	bara (only)	6	66
31.	jobba (work)	3	65
32.	skriva (write)	5	61
33.	svenska (Swedish)	6	60
	dag (day)	4	60
35.	år (year)	6	59
	mig (me)	6	59
	måste (must)	5	59
38.	finnas (exist)	4	58
39.	varje (every)	6	55
	samma (same)	5	55
41.	här (here)	5	52
42.	problem (problem)	5	51
43.	till (to)	6	50
44.	två (two)	5	48
45.	tro (believe)	2	45
46.	vi (we)	3	44
47.	flicka (girl)	5	43
48.	min (my)	6	42
	klockan (clock)	5	42
	spela (play)	5	42
51.	så (so)	3	41
52.	all (all)	5	39
53.	till exempel (for example)	4	38
	säga (say)	3	38
55.	ha (have)	2	37
56.	läsa (read)	5	36
	annan (other)	4	36
	tre (three)	4	36
59.	hjälpa (help)	3	35
	skola (school)	3	35
61.	gång (time, nr of)	5	34
	hemma (at home)	4	34
	ska (will)	4	34
	betala (pay)	3	34
65.	gå (go)	4	33
	oj (oh)	3	33
67.	fru (wife)	4	32
	sex (six)	4	32
	fråga (ask)	4	32
	stanna (stay)	3	32
	att (that, conj.)	2	32
72.	pojke (boy)	6	31
73.	månad (month)	6	29
	absolut (absolutely)	1	29
75.	vänta (wait)	5	28
76.	nu (now)	5	27
	fem (five)	4	27
	stor (big)	4	27
	prata (talk)	3	27
	de (they)	2	27
	okej (OK)	2	27
82.	svår (difficult)	5	26
	många (many)	4	26
83.	titta (look)	4	25
	köra (drive)	3	25
	pengar (money)	3	25
87.	tio (ten)	6	24
	hon (she)	5	24
	maskin (machine)	2	24
	alls (at all)	1	24
92.	fyra (four)	5	22
	idag (today)	5	22
94.	igår (yesterday)	5	21
	liten (small)	4	21
	vill (want)	4	21
	krona (Sw. money)	2	21
	säker (sure)	2	21
99.	man (man)	5	20
	gammal (old)	4	20

One would not expect any major similarities between two so different types of language as the Swedish spoken by immigrants with a very poor command of the language and newspaper-articles of the sort described above, but surprisingly enough, there seem to be some rather striking similarities.

When comparing the list in table 4 with the Top Ten Thousand, one finds that 87 of the most frequent words in Immigrant Swedish are among the 1000 most frequent words in the TTT. Another 8 words in the IS-list occur among the next 3000 most frequent words in the TTT. Looking at the whole IS-material (719 lexical items), one finds that 88% of these words occur in the TTT, i.e. among the 10,000 most common words in written Swedish.

I have also compared the two materials with regard to the words that are most frequent in each word-class. Looking at the 20 most frequently occurring nouns, one finds that 5 words occur in both materials (words meaning 'day', 'year', 'problem', 'time' and 'man'). The rest of these nouns seem mainly to denote persons and time in both of the lists. Among the 20 most frequently occurring verbs, 8 words occur in both materials (meaning 'come', 'be', 'know', 'must', 'have', 'say', 'will' and 'go').

As for the most frequently occurring prepositions, one finds that the two lists are very similar, with words meaning 'on', 'in', 'with', 'to' and 'from' appearing at the top. The same holds true for conjunctions with words meaning 'but', 'and', 'or', 'if', 'that' and 'when'.

On the other hand adjectives, adverbs, pronouns and interjections do not correspond much in the two lists (although there are words that appear in both lists even here). This might be due to the differences between spoken and written language. Interjections, for example, hardly ever occur in newspaper articles, but are very frequent in Immigrant Swedish.

The six informants have roughly the same kind of vocabulary. The two informants who have the smallest vocabulary, Nikolaos and Katarina (about 100 words each), use mainly those words that are the most frequent in Immigrant Swedish, which in turn would mean that most of the words they use are among the most frequent words in Swedish. None of the informants deviate strongly from the others regarding the kind of words he uses. This of course could be due to the fact that they all have more or less the same types of experience in day to day life in Sweden; they meet the same kind of people and use the same kinds of words in their work. Not even the Polish informant deviates much from the others, although in many respects she seems to have come further in the learning process.

Uncommon words

12% of the IS-words do not occur among the 10,000 most frequent words in the TTT. Does that mean that these words are uncommon in Swedish? Looking at some examples, we find words like the following:

Nouns: *gangster* (gangster), *marmelad* (marmalade), *pannkaka* (pancake), *salami* (salami), *vodka* (vodka), *kafeteria* (cafeteria), *orm* (snake), *smuts* (dirt), *grammatik* (grammar), *grekiska* (greek), *påsk* (Easter), *immigrant* (immigrant), *skatt* (tax), *städare* cleaning personnel)
Verbs: *knacka* (knock), *skura* (scrub), *diska* (wash dishes), *tanka* (fill up gasoline), *tystna* (bekome silent)
Adjectives: *turkisk* (Turkish)
Interjections: *hej* (hallo), *varsågod* (please)

These words all are everyday words that rarely occur in newspaper articles of the sort described above, but that possibly are frequent in the spoken Swedish that the informants meet with.

On the learning of words

A material like the one described in this paper is not appropriate if one wishes to study strategies of learning. To do this, one would need either a systematic study over a certain time or some kind of tests. Several observations can however be made, especially if one is interested in looking at the type of lexical elements which result from informal learning. Material of this type can also be used as a basis for a discussion of linguistic and pragmatic factors that may influence the learning of words by learners with low education, a limited experience of studying, and little contact with the target language.

As a starting point for the discussion I shall assume that there is a connection between the occurrence of certain words in the material and the 'learnability' of these words. This would mean that words used by most of the informants, as well as those used in many tokens, tend to be learned earlier than words occurring in a small amount of tokens or those used by only one

or two informants. Such a connection should not be looked upon as being universal, as it may be valid mainly for the informants described here. On the other hand, it may be possible to apply certain aspects of this generalization to learners other than the ones described here.

In the following, three factors which are thought to be important in the process of learning will be discussed: input frequency, complexity and communicative need.

Input frequency

Is the larning of certain lexical items facilitated by the frequency of these words in the target language? I.e. does the learner learn words that he often hears more easily than he learns other, less frequently heard words?

An immigrant who does not go to courses to learn his host country's language immediately upon arrival in the new country, receives, just like the little child learning his first language, an input of the new language mainly in the form of:

a) Language that is addressed directly to him and
b) Language that is spoken around him but not directed to him. This 'overheard' language can be either target-language speakers' conversations which the learner himself does not participate in, or the language that he hears on radio and television.

An adult immigrant may also get linguistic input from reading. Of course this presupposes that the person in question knows how to read, and also that his own alphabet does not differ too much from the one used in the new country.

The Greek and the Latin-Swedish alphabet differ significantly, and none of the Greek informants in this study could be said to be able to read Swedish easily. Katarina did not know the Swedish letters at all, although she knew how to read Greek. No one ever read Swedish newspapers or books. The Polish informant, for whom athe alphabet presented no difficulties, claimed to read a Swedish newspaper every day. It is therefore apparent that language input for the Greek informants almost exclusively consisted of spoken

language, whereas the Polish informant also got written input. But spoken language, even for her, must have been the main source of input.

Ferguson has pointed out (1977:28) that a target-language speaker talking to an immigrant often adapts his language in a way that makes it easy for the learner to comprehend. This 'adaptation' often consists of the use of short sentences, repetition of words to indicate intensity, omission of copula, articles, pronoun subjects and inflectional endings, and so on. This variety of speech has been dubbed 'Foreigner Talk' by Ferguson. 'Foreigner Talk' shows great similarities to 'Baby Talk' (the type of speech which has been observed to be used by adults talking to children, and which is supposed to direct the child's attention to certain linguistic structures and increase understanding).

The other type of input, where the learner does not participate in the communicative act, e.g. TV-programs, conversations between target-language speakers which the learner overhears, written texts, and so on, are undoubtedly at all linguistic levels more complex than the former type.

As mentioned above, the Greek informants seem to have had limited opportunities for linguistic interaction with Swedes in all contexts other than rather official matters. The Polish informant, being married to a Swede, must have had much better possibilities for contact with Swedish-speaking people. As for the 'overheard' conversations, all informants were surrounded by people passing the rooms and corridors where they cleaned, and these people were presumably often engaged in conversations that the informants heard, at least in parts. In the informants' homes, like in most Swedish homes, TV was on during most of the evenings.

The amount of Swedish language input directed at the Greek informants thus seems to have been rather small. 'Overheard' input, however, was much greater, and was more or less equivalent to the amount of 'overheard input' received daily by most native adult Swedes.

To what extent then is the informants' acquisition of vocabulary affected by the quality of input?

Cazden, who (1972:104) discusses how the two different types of input mentioned above affect the child's learning of lexical items, demonstrates that there is a correlation between the frequency of certain linguistic structures in the child's environment and the order of their acquisition. He states however that frequency is not the only important factor. Even a low fre-

quency of a particular linguistic trait may result in learning in cases where the utterances are directed to the child in active interactional situations. Cazden points out that the language that the child listens to on television certainly consists of wellformed sentences, but that it differs from the language used in face-to-face conversation. One of the differences is that language heard on television comes in a highly non-distinctive form, and is simultaneously accompanied by attention-grabbing visual stimuli.

Furthermore, according to Cazden (1972:111), frequency cannot be distinguished from semantic or syntactic complexity. Well-formed sentences that the child can understand and which are uttered in situations where the child is an interactor thus give the most favourable input with regard to learning.

If the same conditions exist for adult learners, this would mean that the language that the immigrant hears at work and on TV and radio influences learning to only a very small extent, while interaction with target language speakers is of considerably greater importance.

We have seen that almost 90% of the words used by the informants also belong to the most frequently occurring words in written, and possibly even in spoken Swedish. Words frequently used in the target language obviously tend to be learned early.

The remaining 10% of the words, which do not occur in the frequency-list of newspaper-language, and probably seldom occur in the input that the learner receives from TV or even in overheard conversations, may nevertheless be relatively frequent in the language that is directed to the learner. These words are often connected to the interactional situations that the immigrant finds himself in.

It thus seems that the informants, even those who have a very restricted vocabulary, have learned two types of words: a) those that are common in all types of Swedish, written as well as spoken, and b) words typical to the situations wherein the informants generally use Swedish.

Complexity

It seems natural to assume that in all kinds of language learning, those items that from the learner's point of view seem simple are learnt before those

which seem more complex. Factors such as similarity to mother tongue words, length of words, semantic complexity, and abstractness and concreteness could be expected to influence the order of acquisition, especially in an informal learning situation.

The words that normally should be the easiest to learn are obviously those which have some kind of resemblance to L_1 words. In this material certain informants use a comparatively large amount of words that have an equivalent in the native language and which sound more or less the same. These words are mainly nouns. About 50% of Katarina's nouns and about 40% of Dimitris's nouns have direct counterparts in Greek. The Polish informant also uses many such words. Most of these words are more or less international, often of Greek origin. Here follow a few examples.

Swedish	Polish	Swedish	Greek
advokat	adwokat	alfabet	alfávito
analfabet (=illiterate)	analfabeta	bebis (=baby)	bébis
		bensin	venzíni
alkoholism	alkoholism	cigarrett	tsigáro
datum	data	demokrati	dimokratía
glob (=globe)	globus	familj	famílja
garage	garaż	kanal	kanáli
grad	grad	klimat	klíma
handla (=buy)	handlować	lexikon	lexikó
studera (=study)	studiować	maffia	mafía

Sometimes such words are used in a form directly transferred from the native language, and sometimes they are adapted phonologically and morphologically to Swedish. The use of such words, common for both the native and the target language, means that the learner does not have to learn new target language lexical items for concepts that have direct equivalents in the native language. He has only to learn that these words exist in the target language as well as in the native language. Corder has called this the 'borrowability' of words.

> 'Successful borrowing', that is when a 'borrowed' item is 'accepted' by the interlocutor as 'well formed' in the target language, may lead to that item being incorporated into the speaker's interlanguage repertoir. This could be regarded as learning. Corder (1981:104)

Kellerman (quoted in Hyltenstam 1978:80) distinguishes language specific elements from language neutral elements in an arbitrary pair of languages. This means that those elements that seem compatible to the learner are language neutral, but that those elements that the learner would not find adequate for use in both languages are language specific. In my material it seems to be the case that learners at an early stage of learning find the language neutral elements and use them. It even seems that the learner sometimes, once having found these words, feels that he knows how to express a certain communicative need. Consequently, he does not bother to learn other words that would be more adequate from the target language point of view, and instead 'over-uses' some of the language neutral words. Dimitris, for example, uses *gangster* whenever he wants to say that a person is in some way unpleasant, and *problem* when he wants to express that someting is 'bad'. Thus he avoids the use not only of synonymous nouns but also of a whole range of adjectives. These words 'block', so to speak, the learning of a whole spectre of Swedish words. (For a further discussion on 'lexical over-use', see Kotsinas 1983a.)

Another factor that probably has great importance for how easy or difficult a word seems to be to the learner is the length of the word. In my material most of the words are short and phonologically uncomplicated. Almost 80% of the 719 words consist of one or two syllables. The rest of the words are mostly uncomplicated verbs consisting of three syllables, or words transferred from the native language. In the list of the 100 most frequent words (table 4), only four words have more than six letters. None has more than seven letters. This of course is connected to the fact that common words in a language tend to be short.

A word that is both frequent in input, short and phonologically simple, may nevertheless be difficult to learn because of its semantic complexity. Herbert Clark claims (1973:29) that if two terms, A and B, are to be learned, A will normally be acquired before B, if B requires all the rules of application of A plus one or more in addition. Clark's discussion concerns child language and deals mainly with the learning of words for time and space, but the assumption may well be applicable to the acquisition of a second language. There is much evidence that words which, according to Clark, are semantically less complex appear in my material in more tokens and are used by more informants than words that are more complex. For example adverbs denoting

position occur more often in the material than adverbs denoting position + direction. Prepositions denoting position + positive direction (to) occur more frequently than prepositions denoting negative direction (from). (For a further discussion see Kotsinas 1982b, 1982c, 1983a.)

Also connected to semantic complexity are over-extensions, frequent in both interlanguage and in child language. Eve Clark (1973:72) has pointed out that children often identify a certain word with the aid of one or two features, instead of with the whole combination of components or features. The child then uses these features to decide whether it is possible in a given context to use a certain word or not. This means that the child makes referential errors because he does not yet know the combinations of features that would otherwise permit him to delimit the categories.

It seems as if the adult learner in certain instances behaves in more or less the same way as a child, in the sense that he first discovers the most important features of a word and then uses this word for all the words within the semantic field where it belongs. This is for example the case with Dimitris, who uses the verb *frdga* (ask) to mean practically all words within the semantic field of 'verbal communication'. Thus *frdga* in certain contexts means 'tell', 'discuss', 'say', 'explain', 'talk' and the like. Likewise he uses *komma* (come) for 'come', 'go', 'travel', and even for 'begin', 'start', 'understand', etc., and *titta* (look) for both 'look' and 'see'. (See Kotsinas 1983a.) Over-extensions in my material seem to occur most frequently among verbs, adverbs and prepositions. This is in contrast to child language, where nouns often are over-extended. Over-extensions occur to different degrees in the speech of different informants.

It is sometimes assumed that a learner prefers concrete words to abstract ones. In my material a great deal of the nouns denote concrete objects and activities, but there are also many abstract nouns. These are often very common words, essential to communication, and they mostly denote time, economy etc. (for example words meaning 'hour', 'month', 'year', 'percentage', 'democracy' etc.). Regarding verbs, there are differences among the informants. Dimitris seems to have concentrated on learning verbs for rather concrete activities, while more abstract verbs that are used rather as form-words are rare in his vocabulary. To replace many of these more abstract verbs, he has worked out an aspectual system, with the aid of the three verbs *komma* (come), *stanna* (stay) and *stoppa* (stop), that shows many resem-

blances to aspectual systems in English-based creole languages. (See Kotsinas 1981a or 1981b.) Maria, on the other hand, has access to many verbs that are close to form-words.

Certain learners may concentrate on learning concrete words as opposed to the more abstract ones because a concrete word seems easier to learn, as it denotes something one can touch or demonstrate. It is easier to ask for the translation of such words by pointing or acting. Another reason for this could be that the learner's idea of language is that it is something which consists of a number of names for things. He wants to know these names, at the same time that he, especially if he is not familiar with grammar or with studies at all, does not consciously understand the value of knowing words that denote abstract concepts, relations, syntactic connections and so on. He may even consider these words to be subtleties that he will not have to learn, because he neither wants nor is able to learn the difficult foreign language perfectly. Finally, it may also be that the speaker, even in his mother tongue, has a rather concrete vocabulary, i.e. he uses certain concretizing strategies to discuss abstract concepts. It is evident that some of the informants, especially Dimitris, have worked out ways to replace more or less abstract expressions by concretizing strategies; using dialogues, numbers, proper names and so on. (See Kotsinas 1982c.)

It would, however, be completely wrong to state that the vocabulary in this material is mainly concrete. The form-words, as well as words denoting the speaker's attitudes and emotions, such as certain adverbs, verbs and nouns, cannot be looked upon as concrete words, and these represent a great part of the informants' vocabulary. It is also evident that different informants follow different patterns when learning new words. It seems for example that Dimitris prefers words for concrete objects and activities, while Maria is more inclined to learn abstract words as well as form-words.

Need for certain words

It is not only the case that the amount or nature of the words that the learner hears affects learning. The learner's need for certain words may be equally important. Three groups of words are distinguishable here: a) words connected with certain situations, where the speaker normally communicates in the

target language, here called situationally important words, b) words that can be used to express the speaker's feelings, attitudes and evaluations, here called emotion words, and c) certain form-words.

As mentioned above, certain words, such as those connected with the speaker's work are likely to be more frequent in the input directed to the speaker than they generally are in the target language. For this reason these words are likely to be acquired early. But it also seems fair to assume that other words, not occurring frequently as input are judged by the learner to be important for him to learn. In this material, there are many nouns and also many verbs which concern subjects that are important to the speaker's everyday life. Examples of such words are words for food, economy, deixis etc., and words like *skatt* (tax), *försäkringskassa* (social security) and *sjukskriva* (report ill). These words are learned because the speaker needs them, and not because they are frequent in input.

Another type of word that the speaker might have a great need for and thus learn early, are words used to express emotions and attitudes, for example interjections (*oj* 'oh'), adverbs (*kanske* 'maybe', *just de* 'exactly', *absolut* 'absolutely', *inte alls* 'not at all'), adjectives (*bra* 'good', *svår* 'difficult', *normal* 'normal') verbs (*måste* 'must', *kan* 'can', *ska* 'shall', *vill* 'want to') and nouns (*vän* 'friend', *problem* 'problem', *fel* 'mistake').

It is evident that the informants' vocabulary does not exclusively consist of words semantically loaded with connection to experiences and needs. A surprisingly great number of form-words are present in the informants' speech. Upon reading certain descriptions of 'immigrants' speech', one may get the impression that form-words are almost totally absent. Meisel, for example, states that articles, prepositions, copula and personal pronouns are 'missing elements' in German Gastarbeiterdeutsch (Meisel 1977:93).

It is indeed true that certain form-word categories, with regard to both tokens and types, are used to a lesser extent in Immigrant Swedish than in normal Native Swedish. This does not mean, however, that form-words are totally missing. Prepositions, for instance, even occur in a higher percentage of tokens in the material than they normally do in spoken Native Swedish. Not even those informants that could be said to know very little Swedish, Katarina and Nikolaos, quite lack auxiliaries, pronouns, articles, conjunctions or prepositions. This indicates that the need of function words is strong and that these words, in spite of the fact that they have a rather vague semantic

meaning, still retain a role so important for communication that they are learnt comparatively early. As Brown & Bellugi point out:

> The meanings that are added ... by functors seem to be nothing less than the basic terms in which we construe reality.
>
> Quoted in Cazden (1972:129)

Marked and unmarked words

It has been suggested that a theory of markedness could explain why certain items or constructions are learned earlier than others. Hyltenstam (1978: 75) suggests the following hypotheses:

H:1) The initial stages of interlanguage are characterized by unmarked categories
H:2) Development towards a given target is achieved from unmarked to marked categories

If in 'unmarked categories' one includes such items that a) are frequent in the target language, b) to the learner seem simple to use, and c) also correspond to fundamental needs of expression, then the hypotheses are supported by my material. It is apparent that situations sometimes occur where these three factors compete, so that words that normally are to be looked upon as complex or uncommon nevertheless are learned because they are needed for a certain purpose. From this follows that one cannot make up lists of words that are frequent and simple and assume that these words are to be taught at early stages of language acquisition. Unless the learner really needs a certain word, it will be difficult for him to learn it, because he will avoid loading his memory with 'unnecessary' stuff.

On the other hand, when the learner has a strong need to communicate, he will use almost any lexical material to fulfil this need.

Strategies of learning

As previously stated, the material presented here is not suitable for investigating how the learning of vocabulary is technically achieved, i.e. which strategies a learner more or less consciously uses to apprehend or memorize lexical items. There are, however, a few suggestions that can be made concerning strategies of learning.

Corder proposes that the learner may sometimes acquire a structure as a holophrase and produce it on the appropriate occasion and without knowing the rules for generating that structure (Corder 1975:127).

In my material there are quite a few phrases and constructions that may well have been learned as 'whole chunks' without the learner being conscious of the function or meaning of the different elements, or without his being able to use them in other contexts. Some of these chunks are idiomatic phrases. Others are groups of words that often occur together in the same kind of contexts, without necessarily being looked upon as lexicalized items. An example of the first type is: *det spelar ingen roll* (it doesn't matter, lit. it does not play any role). An example fo the second type is: *Hur mår du? Bra*? (How are you, how do you feel? Fine?). Dimitris often uses the second phrase, always in the same way, without any variation and, it seems, without expecting any answer. It appears that he uses it as some sort of polite greeting, not being able to analyze the different parts of it. He never uses *hur* (how) or *mår* (feel) in other contexts. Takis very often uses *de e* (it is) or *den e* (it is) at the beginning of a sentence. He uses the phrase as if it were some kind of starting-point for the sentence without any particular semantic content of its own. Likewise *är det*? (is it?) seems to be a question marker rather than verb-subject of the sentence. Other more or less stereotyped combinations of words are *här inne* (in here), *där nere* (down there), *gå ut* (go out), *komma in* (come in), *gå upp* (go up), *go ner* (go down), *du vet* (you know), *förstår du mig*? (do you understand me?). The words *inne, nere, ut, in, upp, ner* are never used in other contexts.

Another strategy of learning that can be discerned in the material is what I call copying, viz. certain words or phrases are used by a certain informant only directly after the same word or phrase has been used by one of the other informants or by the teacher. Dimitris for example, normally uses *komma* (come) to denote all kinds of motion ('come', 'go', 'travel', 'run' etc.). In a

total of 109 examples, he uses *komma* in all but four instances (where he uses *gå* (go) instead). He uses *gå* correctly, but does so directly after he has just heard someone else use the verb. Only twice does he use the verb *gå* in 'free production'. This would mean that *gå* exists as a more or less passive bit of knowledge in his mind, but that he does not dare to use it unless he has heard the word being used by someone else in a context where he understands it. The two cases where he uses the word without its being present in the situation could indicate that he is slowly beginning to risk using this word instead of the more familiar *komma*, which he normally uses also when he means 'go, 'leave'.

It seems plausible to assume that at least some lexical items are introduced into the speaker's vocabulary first as parts of more or less stereotyped chunks, and that they later become 'emancipated' when the learner starts to use them out of their 'usual' context.

Other factors that affect the learning of lexical items

The factors discussed above concerning the acquisition of vocabulary are all more or less connected to the linguistic situation. It is of course clear that non-linguistic factors also steer the process of acquisition. These factors possibly account for the well-known fact that some learners go on learning until they reach a competence that is very near to the native speaker's, while others seem to cease at a certain stage and fossilize there. Naturally factors such as motivation, education, contact with target language speakers, attitudes to the new country, personality etc. are of great importance. These factors are not directly investigated in this study but they can be inferred. Conversely, age and length of stay in the host-country do not seem to have as great an importance as the factors mentioned above. Maria was 46 years old when she immigrated, but has a much larger vocabulary than Nikolaos, who immigrated at the age 23 and who, like Maria, had spent 3 years in Sweden. Maria's vocabulary seems to be about the same size as Dimitris's, but in many respects her language is closer to native Swedish than Dimitris's, even though he was 35 years old upon immigration, and had been in Sweden three times as long. One can also postulate that Maria will go on learning at all linguistic levels, while Dimitris will more or less stay at the stage where he is now, with the exception that he will learn a few more lexical items.

Summary

In this paper, the kinds of words a learner acquires in an informal learning situation have been discussed. Simple and common words form the main part of the vocabulary of the informants. The following factors operate in the process of learning:

1. Frequency of input words
   a) Frequency of words in the target language
   b) Frequency of words in situations where the learner usually interacts linguistically

2. Complexity
   a) Similarity to native language words
   b) Length of words
   c) Semantic compexity
   d) Degree of concreteness or abstractness

3. Communicative need of certain words
   a) Words connected to the situation where the learner uses the target language
   b) Words expressing emotions and attitudes
   c) Form-words

Although a material such as this, consisting as it does of spontaneous utterances, cannot be used to investigate all strategies of learning, it is possible to draw some inferences concerning these strategies. The following strategies are discussed and exemplified:

1. The learning of 'whole chunks'
2. The learning of words occurring in certain contexts
3. Copying

REFERENCES

Allén, S. (1972), *Tiotusen i topp*. Stockholm: Almqvist & Wiksell.

Cazden, C.B. (1972), *Child Language and Education*. New York et al.: Holt, Rinehart and Winston, Inc.

Clark, E.V. (1973), What's in a word? On the child's acquisition of semantics in his first language. Moore, T.E. (ed.), *Cognitive Development and the Acquisition of Language*. New York & London: Academic Press, 65-103.

Clark, H.H. (1973), Space, time, semantics and the child. Moore T.E. (ed.), *Cognitive Development and the Acquisition of Language*. New York & London: Academic Press, 27-63.

Corder, S.P. (1975), Error analysis. *The Edinburgh course in applied linguistics*, vol. 2. London: Oxford University Press, 122-131.

Corder, S.P. (1981), *Error analysis and interlanguage*. London: Oxford University Press.

Einarsson, J. (1978), *Talad och skriven svenska*. Lundastudier i nordisk språkvetenskap. Lund: Walter Ekstrands bokförlag.

Ferguson, Ch.A. (1977), Simplified Registers, Broken Language and Gastarbeiterdeutsch. Molony, C.H., Zobl, W. & Stölting, W. (eds.), *Deutsch im Kontakt mit anderen Sprachen – German in Contact with Other Languages*. Kronberg/Ts: Scriptor, 25-39.

Hyltenstam, K. (1978), A framework for the study of interlanguage continua. *Working Papers*, Phonetics Laboratory, Department of General Linguistics. Lund: Lund University, 16, 65-86.

Kotsinas, U.-B. (1981a), Kommer och predikatet GO. Funderingar kring tempus och aspekt i invandrarsvenska. Ejerhed, E. & Henrysson, I. (eds), *Tvåspråkighet*. Umeå universitet, Acta Universitatis Umensis, Umeå Studies in the Humanities 36, 201-213.

Kotsinas, U.-B. (1981b), Come and GO: Tense and Aspect in Immigrant Swedish. Dittmar, N. & Königer, P. (eds), *Linguistische Arbeiten und Berichte*, Heft 16. Berlin: Fachbereich Germanistik der Freien Universität, Berlin, 112-129.

Kotsinas, U.-B. (1982a), Omtagning – en textstrategi i invandrarspråk. Koch, W., Platzack, Ch. & Tottie, G. (eds), *Textstrategier i tal och skrift*. Stockholm: Almqvist & Wiksell international, 191-212.

Kotsinas, U.-B. (1982b), Kommer en papper på kyrka. Lokativa prepositioner i invandrarsvenska. *Svenskans beskrivning 13*. Helsinki: Meddelanden från institutionen för nordiska språk och nordisk litteratur vid Helsingfors universitet, Serie B nr 6 1982.

Kotsinas, U.-B. (1982c), *Svenska svårt. Några invandrares svenska talspråk. Ordförrådet*. Stockholm: Institutionen för nordiska språk vid Stockholms universitet, MINS 10.

Kotsinas, U.-B. (1983a), Semantic over-extension and lexical over-use in Immigrant Swedish. *Scandinavian working Papers on Bilingualism*, Stockholm: University of Stockholm, Institute of Linguistics, Department of Research on Bilingualism (To appear).

Kotsinas, U.-B. (1983b), Repetitions in Immigrant Swedish. *Indian Journal of Applied Linguistics*. Syracuse, New York. (To appear).

Meisel, J. (1977), Linguistic Simplification: A Study of Immigrant Workers' Speech and Foreigner Talk. Corder, S.P. & Roulet, E. (eds), *The Notions of Simplification, Interlanguages, and Pidgins and their Relation to Second Langauge Pedagogy*. Genève: Droz, 88-98.